

Performance of the CORE-10 and YP-CORE measures in a sample of youth engaging with a community mental health service

AILEEN O'REILLY,¹ NICHOLAS PEIPER,² LYNSEY O'KEEFFE,¹ ROBERT ILLBACK^{1,3}
& RICHARD CLAYTON⁴

1 Headstrong – The National Centre for Youth Mental Health, Dublin, Ireland

2 Behavioral and Urban Health Program, RTI International, Research Triangle Park, NC USA

3 REACH of Louisville, Inc., Louisville, KY USA

4 Department of Health Behavior, University of Kentucky, Lexington, KY USA

Key words

Clinical Outcomes in Routine Evaluation (CORE) measures, psychological distress, early intervention, psychometrics, epidemiology

Correspondence

Aileen O'Reilly, Headstrong – The National Centre for Youth Mental Health, 16 Westland Square, Pearse Street, Dublin 2, Ireland.
Telephone 00 353 1 4727 058
Email: aileen.oreilly@headstrong.ie

Received 1 July 2015;
revised 16 November 2015;
accepted 17 November 2015

Abstract

This article assesses the performance and psychometric properties of two versions of the Clinical Outcomes in Routine Evaluation (CORE) measures that assess psychological distress: the Young Person's CORE (YP-CORE) for 11–16 year olds and the CORE-10 for those 17 or older. The sample comprised 1592 young people aged 12–25 who completed the YP-CORE and CORE-10 during their initial engagement with an early intervention service. Total and average scores were examined for both measures. Gender and age differences were evaluated using *t*-tests and analysis of variance. The factor structures were assessed with principal axis and confirmatory factor analyses. Multigroup confirmatory factor analyses were then employed to evaluate measurement invariance across age and gender. Analyses were supportive of the CORE measures as reliable instruments to assess distress in 12–25 year olds. Based upon eigenvalues in combination with the comparative fit index, the Tucker–Lewis Index, and the root-mean-square error of approximation, both measures were unidimensional. Analysis indicated the factor structure, loadings, item thresholds, and residuals were invariant across age and gender, although partial support for strict invariance was found for gender among 12–16 year olds. Results are compared to previous studies and discussed in the context of program planning, service delivery, and evaluation. Copyright © 2016 John Wiley & Sons, Ltd.

Introduction

Mental health practitioners are increasingly being encouraged to adopt evidence-based practices, and the collection of outcome data has been highlighted as a priority for mental health services. The Clinical Outcomes in Routine Evaluation Outcome Measure (CORE-OM) is one outcome measure that has been widely used to evaluate psychological distress during treatment (Barkham *et al.*, 2006). The CORE-OM consists of 34 items which cover four domains: *subjective well-being*, which refers to an individual's quality of life; *symptoms*, which comprises item clusters of depression, anxiety, trauma and physical correlates of psychological health; *functioning*, which refers to general functioning in daily life and relationships; and *risk*, which covers self-harm, suicidal ideation, and risk to self/others (Skre *et al.*, 2013). The CORE-OM has been widely used in the evaluation of counseling and psychological therapies in the United Kingdom (UK). Furthermore, it has been validated for use with samples from the general population, primary and secondary care, and older adults (Barkham *et al.*, 2005a; Barkham *et al.*, 2005b; Connell *et al.*, 2007; Evans *et al.*, 2003; Mellor-Clark *et al.*, 2001).

Additional measures of psychological distress have been developed by the authors of the CORE-OM including the CORE Short Forms A&B for session-by-session use in research settings (Cahill *et al.*, 2006), a version for the general population (GP-CORE) (Sinclair *et al.*, 2005), and a questionnaire for individuals with learning disabilities (LD-CORE; Marshall and Willoughby-Booth, 2007). The CORE-10 has also been developed for use where time and resource demands render administration of the CORE-OM infeasible (Barkham *et al.*, 2013). Data used for the development of the CORE-10 consisted of approximately 6000 clients from 33 primary care services in the UK. Similar to the CORE-OM, items on this measure comprise questions about subjective well-being, symptoms of depression and anxiety, general functioning and risk. The measure has been manualized (Connell and Barkham, 2007) and has been shown to have strong psychometric properties. For example, alpha levels of 0.89 (confidence interval [CI] 0.84 to 0.92) and 0.89 (CI 0.80 to 0.94) have been reported for female and male samples, respectively (Barkham *et al.*, 2013).

Emerging research has provided support for redesigning mental health services for young people aged 12–25 years, in recognition of the fact that mental ill health is the most important health issue facing this age group worldwide (McGorry *et al.*, 2013; McGorry *et al.*, 2014). Recognizing the importance of this growing body of evidence, another version of the CORE, the YP-CORE, has been validated by the developers for young people aged 11–16 years (Twigg *et al.*,

2009). As with the CORE-10, there are 10 items on the YP-CORE that measure subjective well-being, commonly experienced problems or symptoms, risk, and life/social functioning (Twigg *et al.*, 2009; Twigg and McInnes, 2010). These items were initially selected for inclusion in the YP-CORE by a researcher and mental health professionals, and the questionnaire was subsequently piloted with samples of young people in clinical and non-clinical settings. Normative data for the YP-CORE are available for young people engaging with psychological interventions in the UK and a limited non-clinical group of young people in educational settings (Twigg *et al.*, 2009). Analysis of the psychometric properties of the YP-CORE has revealed that the measure has acceptable internal reliability (Cronbach's alpha of 0.85, CI 0.77 to 0.88), a discernible factor structure, and is sensitive to meaningful change (Twigg *et al.*, 2009; Twigg and McInnes, 2010).

Although research on the use of the CORE-10 and YP-CORE is promising, their psychometric properties have not been assessed in a community-based sample of young people. Further examination of the performance of the YP-CORE is warranted, given that this is a relatively new measure. The present study evaluates the performance of the CORE-10 and YP-CORE among a sample of young people aged 12–25 years engaging with the Jigsaw early intervention community mental health service in the Republic of Ireland. Jigsaw was developed by Headstrong: The National Centre for Youth Mental Health and provides support to young people across Ireland in close collaboration with statutory, voluntary and community mental health and related services (O'Keeffe *et al.*, 2015; Peiper *et al.*, 2015; O'Reilly *et al.*, 2015). The CORE measures are utilized in Jigsaw to estimate the prevalence of distress among young people and to evaluate the effectiveness of the service.

Methods

Sample

Participants were 1592 young people (aged 12–25 years) who engaged with Jigsaw for a brief intervention between September 1, 2013 and December 31, 2014. A brief intervention consisted of goal-focused face-to-face engagements with a young person, and typically lasted between one to six sessions. The average number of brief intervention sessions for young people engaging with Jigsaw was 4.4 sessions. The current sample consisted of 700 young people who completed the YP-CORE (12–16 year olds) and 892 young people who completed the CORE-10 (17–25 year olds) during their first brief intervention session in Jigsaw.

Procedure

Information about young people who engaged with Jigsaw was captured using the online Jigsaw Data System (JDS). All young people (and parents/guardians, where appropriate) consented to their details being recorded on the JDS and anonymously analyzed for research and evaluation purposes. For the purposes of this study, data accessed by the research team consisted of participants' age, gender and CORE responses.

Ethical approval for this research was granted by the Institutional Review Board of the University of Kentucky Office of Research Integrity, due to the authors' affiliation with this University. The study was also carried out in accordance with local clinical governance arrangements. The ethical review carried out was based on legal and ethical standards consistent with Irish standards, including parental consent for those under age 18 and standard components of informed consent.

Measures

Items on the CORE measures are scored from zero to four, with higher scores indicating greater levels of distress. The total scale score is calculated by summing each response, dividing by the number of items answered and multiplying by 10, thus the range is zero to 40. Although a clinically validated cutoff has yet to be determined for the YP-CORE, a score of 11 or above on the CORE-10 indicates clinically significant distress. In addition, scores on the CORE-10 can be divided into categories of distress: Healthy (0–5), low (6–10), mild (11–14), moderate (15–19), moderate-to-severe (20–24), and severe (25 and above). These cutoffs were established by asking a large sample of the UK population to complete the questionnaire and comparing their scores statistically with those for large samples of clients in therapy (Connell and Barkham, 2007).

Data analysis

To examine levels of psychological distress, the prevalence and cumulative prevalence for each total score were computed for the YP-CORE ($n=700$) and CORE-10 ($n=892$). Descriptive statistics, including means (M values), standard deviations (SD values), and 95% confidence intervals (CI) were calculated by age, gender, and overall. For the CORE-10, the documented cutoffs were used to examine the prevalence of clinically significant distress and distribution of distress categories. To determine significant population-based differences by age and gender, independent samples t -tests and analysis of variance (ANOVA) were carried out for both measures. Chi-square statistics

were also employed to determine age and gender differences for the CORE-10 cutoffs.

To examine the factor structure of the CORE measures, polychoric correlations were first calculated for the CORE-10 and YP-CORE items (Kolenikov and Angeles, 2004; StataCorp, 2012). Next, principal axis factor analysis was carried out using the generated 10-by-10 matrices of polychoric correlations, and internal consistency was evaluated with Cronbach's alpha. Confirmatory factor analysis (CFA) was then executed to corroborate the dimensionality of both measures using the following indices: for the comparative fit index (CFI) and the Tucker–Lewis Index (TLI) > 0.90 was considered adequate and > 0.95 was very good; and for the root mean square error of approximation (RMSEA) < 0.08 was considered adequate and < 0.05 was very good (Hu and Bentler, 1999). Multigroup CFAs were also performed to determine the equivalence of the CORE factor structures across age and gender. This involved a series of constraints imposed at four successive levels of stringency (Gregorich, 2006): a configural invariance model (parameters freely estimated), a metric invariance model (factor loadings constrained to be equal across groups), a scalar invariance model (factor loadings and item thresholds constrained), and a strict invariance model (factor loadings, item thresholds, and residuals constrained). Finally, measurement invariance was evaluated by change in fit in comparison to the previous model: change in CFI (CFI) < 0.01 and change in TLI (TLI) < 0.02 as supportive of the more stringent model (Cheung and Rensvold, 2002).

All analyses were performed with Stata 12.1 and a two-sided p -value of < 0.05 was considered statistically significant.

Results

Distribution and descriptive statistics

For the YP-CORE, approximately 47% of 12 to 16 year olds scored 0–17; and for the CORE-10, 47% of 17 to 25 year olds scored 0–19. The descriptive statistics for each measure are shown in Table 1. For 12 to 16 year olds, the average YP-CORE score was approximately 18 ($M=18.3$, $SD=7.6$, 95% $CI=17.7$ – 18.9). Significant age differences were found, $t(698)=-5.16$, $p<0.001$, with 12–13 year olds having lower scores ($M=15$, $SD=7.5$, 95% $CI=13.6$ – 16.4) than 14–16 year olds ($M=18.9$, $SD=7.6$, 95% $CI=18.3$ – 19.5). There was also a significant gender effect, $t(698)=6.33$, $p<0.001$, with females having higher scores ($M=19.7$, $SD=7.4$, 95% $CI=19.0$ – 20.4) than males ($M=16.7$, $SD=7.4$, 95% $CI=15.9$ – 17.6). For 17 to 25 year olds, the average CORE-10 score was about 20 ($M=19.6$, $SD=6.6$, 95% $CI=19.2$ – 20.1). A significant gender effect was demonstrated, $t(890)=4.01$, $p<0.001$,

Table 1. Descriptive statistics from the YP-CORE and CORE-10 items

	Alpha	Mean (SD)	95% CI	<i>t/F</i> (<i>p</i> -value)
YP-CORE (<i>n</i> =700)	0.81	18.3 (7.6)	17.7–18.9	
Age				–5.16 (0.001)
12–13	0.80	15.0 (7.6)	13.6–16.4	
14–16	0.81	18.9 (7.5)	18.3–19.5	
Gender				6.33 (0.001)
Female	0.81	19.7 (7.4)	19.0–20.4	
Male	0.79	16.7 (7.4)	15.9–17.6	
CORE-10 (<i>n</i> =892)	0.77	19.6 (6.6)	19.2–20.1	
Age				1.65 (0.193)
17–19	0.78	19.3 (6.8)	18.7–19.9	
20–22	0.74	20.2 (6.3)	19.5–20.9	
23–25	0.76	19.6 (6.4)	18.5–20.7	
Gender				4.01 (0.001)
Female	0.76	20.3 (6.4)	19.8–20.8	
Male	0.77	18.5 (6.8)	17.8–19.2	

Abbreviation: SD, standard deviation; CI, confidence interval.

with females having higher scores ($M=20.3$, $SD=6.4$, 95% $CI=19.8$ – 20.9) than males ($M=18.5$, $SD=6.8$, 95% $CI=18.8$ – 19.3). No significant effects were found by age, $F(2, 889)=1.65$, $p=0.19$.

While the YP-CORE has yet to develop clinically validated cutoffs, roughly 91% of 17 to 25 year olds screened positive for clinically significant distress on the CORE-10 (i.e. score of 11 or above). Approximately 12.2% of participants' scores fell in the mild range, 25.6% were moderate, 28.9% were moderate-to-severe, and 23.9% were severe. In addition, a significantly higher proportion of females were in the moderate-to-severe and severe categories, $\chi^2(4)=13.73$, $p=0.008$. No other significant differences were found using the dichotomous cutoff or among the severity categories.

Internal consistency and factor structure

Cronbach's alpha for the YP-CORE and CORE-10 was 0.81 and 0.77, respectively. Across age groups and gender, Cronbach's alpha ranged from 0.74 to 0.78 for the YP-CORE and 0.79 to 0.81 for the CORE-10 (Table 1). The items on the YP-CORE had polychoric correlations that ranged from 0.08 to 0.67. In the principal axis factor analysis of YP-CORE items using oblique rotation (Table 2), the first factor (eigenvalue = 3.42) comprised eight items with factor loadings above 0.4, including four items from the *problems* domain, two from the *functioning* domain, one from the *subjective well-being* domain, and one from the *risk to self* domain. The second factor (eigenvalue = 0.73) had two

items with factor loadings above 0.4 that mainly captured general functioning and social relationships. One item *I haven't felt like talking to anyone* loaded on both factors (0.49 for factor 1 and 0.47 for factor 2). The correlation between factors was 0.33.

For the CORE-10, the polychoric correlations ranged from 0.12 to 0.73. In the rotated principal axis factor analysis, the first factor (eigenvalue = 3.2) comprised five items, including two from *problems*, one from general *functioning*, one from harm to *self*, and one from *trauma*. The remaining five items had marginal loadings that ranged from 0.3 to 0.39. The second factor (eigenvalue = 0.35) had two items that mainly captured the anxious symptoms from the *problems* domain. One item *I have difficulty getting to sleep or staying asleep* moderately loaded on this factor (0.37). The correlation between factors was 0.49.

Confirmatory factor analyses

Because the eigenvalues from the exploratory factor analysis (EFA) suggested a one-factor structure (i.e. first eigenvalue above one and second value below one) for both measures, one-factor models were specified in the CFAs (Table 3). The one-factor solution had strong absolute fit for the YP-CORE, $\chi^2(29)=43.04$, CFI = 0.993, TLI = 0.988, RMSEA = 0.026 and CORE-10, $\chi^2(29)=43.15$, CFI = 0.992, TLI = 0.989, RMSEA = 0.022. In the multigroup CFAs, all of the invariance models by age and gender had acceptable absolute fit for the

Table 2. Principal axis factor analysis of the polychoric correlation matrix of CORE items

		One-factor	Two-factor	
CORE items	Mean (SD)	Factor 1	Factor 1	Factor 2
<i>YP-CORE</i>				
Eigenvalue			3.42	0.73
Edgy or nervous	1.91 (1.19)	0.62	0.65	0.03
Talking to others	1.83 (1.23)	0.62	0.49	0.47
Able to cope	1.97 (1.19)	0.55	0.51	0.21
Thoughts of self-injury	0.65 (1.07)	0.61	0.55	0.27
Able to ask for help	1.58 (1.35)	0.26	0.11	0.48
Distressed thoughts	2.23 (1.23)	0.76	0.77	0.10
Problems too much	1.96 (1.27)	0.80	0.81	0.12
Sleep problems	1.88 (1.53)	0.54	0.52	0.16
Felt unhappy	2.19 (1.17)	0.81	0.76	0.27
Done all things	2.10 (1.27)	0.37	0.32	0.19
<i>CORE-10</i>				
Eigenvalue			3.20	0.35
Anxious or nervous	2.71 (1.03)	0.58	0.33	0.59
Someone to turn to	1.64 (1.19)	0.30	0.30	0.08
Able to cope	2.31 (1.04)	0.49	0.45	0.20
Talking to others too much	1.91 (1.18)	0.41	0.34	0.23
Panic or terror	1.69 (1.25)	0.60	0.39	0.52
Suicide plan	0.36 (0.79)	0.54	0.59	0.04
Sleep problems	2.43 (1.34)	0.46	0.32	0.37
Despair or hopeless	2.03 (1.24)	0.82	0.79	0.25
Felt unhappy	2.65 (1.01)	0.78	0.76	0.25
Unwanted images/memories	1.90 (1.40)	0.48	0.40	0.25

Abbreviation: SD, standard deviation.

YP-CORE. The change in fit was also demonstrated across age and gender for all levels of invariance, with the exception of strict invariance that significantly degraded from the scalar model for gender (CFI=0.027, TLI=0.022). The absolute fit for the strict model remained acceptable, however, and RMSEA also fell into the CI for the scalar model, offering partial support of strict invariance. In the models for the CORE-10 across age and gender, absolute and change in fit values were acceptable. These results indicate that the one-factor structure, factor loadings, item thresholds, and residuals are all invariant across age and gender for both CORE measures.

Discussion

The purpose of this study was to evaluate the psychometric properties of the YP-CORE and CORE-10 measures and to validate these measures for use with young people

in community mental health settings. Overall, the results suggest both measures perform well as short screening scales for psychological distress within the context of early intervention community services targeting young people 12–25 years old. The distribution of YP-CORE scores showed approximately half of 12 to 16 year olds scored 0–17. Similarly, for the CORE-10, nearly 50% of 17 to 25 year olds scored 0–19.

In addition, a significant proportion of young people self-reported high levels of psychological distress during their first brief intervention session in Jigsaw. While these levels of distress are higher than those found in population-based samples (Connell and Barkham, 2007), comparable findings have been demonstrated in a variety of clinical samples in the UK and Australia, including primary care, general practitioner (GP) practice, early intervention services, and youth counseling services (Barkham *et al.*, 2013; Rickwood *et al.*, 2014; Twigg *et al.*, 2009). Moreover, the significant age- and gender-based differences among Irish youth are broadly

Table 3. Multigroup confirmatory factor analyses for measurement invariance

Model	$\chi^2_{(df)}$	CFI	TLI	RMSEA	$\Delta\chi^2$	Δ CFI	Δ TLI
YP-CORE	43.04 ₍₂₉₎	0.993	0.988	0.026			
Age							
Configural	75.17 ₍₅₈₎	0.991	0.985	0.029			
Metric	86.04 ₍₆₇₎	0.989	0.986	0.028	10.87	0.002	0.001
Scalar	98.97 ₍₇₆₎	0.987	0.985	0.029	12.93	0.002	0.001
Strict	117.69 ₍₉₂₎	0.986	0.986	0.028	18.72	0.001	0.001
Gender							
Configural	71.30 ₍₅₈₎	0.992	0.988	0.026			
Metric	86.75 ₍₆₇₎	0.989	0.985	0.029	15.45	0.003	0.003
Scalar	106.71 ₍₇₆₎	0.983	0.979	0.034	19.96	0.006	0.006
Strict	169.45 ₍₉₂₎	0.956	0.957	0.049	62.74	0.027	0.022
CORE-10	43.15 ₍₃₀₎	0.992	0.989	0.022			
Age							
Configural	125.20 ₍₉₆₎	0.983	0.976	0.032			
Metric	146.30 ₍₁₁₄₎	0.981	0.978	0.031	21.10	0.002	0.002
Scalar	172.01 ₍₁₃₂₎	0.977	0.977	0.032	25.71	0.004	0.001
Strict	206.52 ₍₁₅₈₎	0.972	0.976	0.032	34.51	0.005	0.001
Gender							
Configural	70.79 ₍₆₂₎	0.995	0.992	0.018			
Metric	78.01 ₍₇₁₎	0.996	0.995	0.015	7.22	0.001	0.003
Scalar	95.59 ₍₈₀₎	0.991	0.990	0.021	17.58	0.005	0.005
Strict	115.77 ₍₉₄₎	0.987	0.988	0.023	20.18	0.004	0.002

Abbreviation: $\chi^2_{(df)}$, chi-square (degrees of freedom); CFI, comparative fit index; TLI, Tucker–Lewis Index; RMSEA, root mean square of error of approximation; Δ , change.

consistent with clinical and population-based studies (Kessler *et al.*, 2012; Merikangas *et al.*, 2010; Merikangas *et al.*, 2009).

With regards to psychometric properties, both CORE measures were shown to have good reliability across age groups and gender, which accords with previous research (Barkham *et al.*, 2013; Twigg *et al.*, 2009). Although limited information has been published to date about the factor structure of the YP-CORE and CORE-10, this study indicates that the two measures have a single factor structure. This conflicts with findings from the developers of the YP-CORE, who suggested a two-factor solution for the YP-CORE. These results may stem from the different methodological approaches employed. For example, Twigg *et al.* (2009) utilized principal components analysis (PCA) to derive a two-factor solution for the YP-CORE, while this study employed principal axis factor analysis to explore dimensionality. This distinction has important implications for the eigenvalues as PCA does not discriminate between shared and unique variance, leading to inflated values (Costello and Osborne, 2005; Gorsuch, 1997). Indeed, we yielded eigenvalues of 4.36 and 1.09 for the YP-CORE with PCA of the polychoric correlation matrix, which were virtually identical to 4.43 and 1.08

found by Twigg *et al.* (2009). Consequently, PCA tends to be more appropriate for data reduction situations, while principal axis analysis tends to be a more practical method to determine dimensionality (Costello and Osborne, 2005). Considering the EFA and CFA findings together, it appears the one-factor solutions for both CORE measures effectively capture the interrelated aspects of psychological distress, including physical symptoms, relationships, general functioning, risk to self and well-being.

Multigroup CFAs carried out in this study indicated structural equivalence across gender and age groups on both the YP-CORE and CORE-10. Furthermore, the invariance demonstrated in the more stringent models lends additional evidence to the measurement equivalence of the CORE measures. This is crucial for making population-based comparisons, implying that the difference in mean scores on the YP-CORE and CORE-10 represent true gender and age differences in psychological distress as opposed to systematic item response patterns influenced by group membership. The partial attainment of strict invariance for gender on the YP-CORE, however, suggests meaningful comparisons of observed variance estimates may not be fully defensible. This

is consistent with a multitude of studies that argue in favor of demonstrating equivalence of residual variances in addition to factor loadings and item thresholds (DeShon, 2004; Wu *et al.*, 2007), although others argue strict invariance may not be a necessary or practical condition (Gregorich, 2006; Vandenberg and Lance, 2000). Nonetheless, the degradation in fit of the strict invariance model compared to the strong model suggests that there are important gender differences in variances of the individual factors that are unaccounted for by the latent factor of psychological distress (Kramer *et al.*, 2008). Although analysis of the specific risk factors for psychological distress that account for gender differences in variance was beyond the scope of this study, it is likely that varying prevalence rates, mean levels, and rates of exposure account for these differences. Based upon these findings and the ongoing debate regarding strict invariance, thorough comparisons of variance estimates from multiple samples will be required to determine the necessity for gender-normed versions of the YP-CORE.

Several limitations warrant further research. While the present study evaluated the construct validity of two CORE measures with regard to dimensionality and measurement invariance, we were unable to directly assess convergent and discriminant validity. Furthermore, with regard to the CORE-10, the adult population in this study was limited to 17–25 year olds. Given the incidence and variability of psychological distress across the lifespan (Burstein *et al.*, 2012; Jorm *et al.*, 2005), adult studies of the CORE-10 are required to evaluate measurement invariance across a broader age spectrum. Similarly, the lack of verified classification functions for the YP-CORE precluded the ability to comprehensively evaluate clinically significant change among the full sample of young people. While such values exist for the CORE-10, performing such analyses among 17–25 year olds would have left a significant gap with regard to clinical outcomes among the younger individuals engaging in brief interventions. Furthermore, the overarching goal of this study was to determine structural and measurement equivalence across key demographic groups. For instance, non-equivalence of the CORE measures would have indicated that age- or gender-related differences in the prevalence of clinically significant distress would not have been a function of mean-level differences in the underlying factor of non-specific distress. Thus, deriving an accurate and invariant latent structure represents a crucial first step in determining whether measures of psychological distress may validly make comparisons across key demographic groups and predict outcomes of interest. As this study indeed demonstrated measurement invariance for the two CORE measures, upcoming studies will more thoroughly evaluate clinical outcomes among relevant subpopulations of Irish youth.

Longitudinal invariance and treatment effects also warrant evaluation as this study employed a cross-sectional design with young people at intake of an early intervention service (Drapeau *et al.*, 2010). Similarly, the impact of other socio-demographic characteristics such as ethnicity on levels of psychological distress bares critical importance. According to the 2011 Census, approximately 12% of the total population in Ireland identified their ethnic or cultural background as non-Irish and roughly 750,000 individuals indicated they were born outside of Ireland (Central Statistics Office, 2012). Between 2006 and 2011, the Traveler community increased by 32%, with the numbers of Traveler youth increasing from 10,929 to 14,245 (Central Statistics Office, 2012). As these populations disproportionately experience marginalization, stigma, poverty and discrimination (Mental Health Reform, 2014), it remains imperative to address methodological issues related to the perception and expression of mental health. At present, only a small number of Irish Travelers and young people from different ethnic backgrounds (approximately 1% and 7%, respectively) have engaged with the Jigsaw service. However, as these numbers increase, studies will be necessary to determine measurement invariance of the CORE measures across ethnic and cultural groups.

Our study is the first to examine the psychometric properties of the CORE measures among a large sample of young people in a community mental health setting. The results from this study suggest the CORE-10 and YP-CORE measures have strong psychometric properties and that the measures are invariant across gender and age. As the numbers of young people engaging with Jigsaw continues to increase, use of the YP-CORE and CORE-10 appears appropriate for assessment, outcome evaluation, and epidemiologic purposes. However, clinical calibration studies that optimize the YP-CORE's classification functions are necessary to determine cutoff points for distress categories, generate prevalence estimates, and inform the services provided to young people aged 12–16 years. The adoption of these measures in other community mental health settings may further elucidate on the generalizability of our findings and provide additional opportunities to evaluate scale performance.

Acknowledgements/Acknowledgments

The work of Drs Illback, O'Reilly and O'Keeffe at Headstrong was supported by philanthropic funding from the One Foundation and Atlantic Philanthropies. Drs Peiper and Illback were supported by the US Substance Abuse and Mental Health Services Administration (Grant SP019436-01-2). Dr Clayton was on sabbatical leave from the University of Kentucky, College of Public Health and worked as a Senior Research Scientist at Headstrong in 2012.

Declaration of interest statement

The authors have no competing interests.

References

- Barkham M., Bewick B., Mullin T., Gilbody S., Connell J., Cahill J., Mellor-Clark J., Richards D., Unsworth G., Evans C. (2013) The CORE-10: a short measure of psychological distress for routine use in the psychological therapies. *Counselling and Psychotherapy Research*, **13**(1), 3–13. DOI:10.1080/14733145.2012.729069.
- Barkham M., Culverwell A., Spindler K., Twigg E. (2005a) The CORE-OM in an older adult population: psychometric status, acceptability, and feasibility. *Aging & Mental Health*, **9**(3), 235–245. DOI:10.1080/13607860500090052.
- Barkham M., Gilbert N., Connell J., Marshall C., Twigg E. (2005b) Suitability and utility of the CORE-OM and CORE-A for assessing severity of presenting problems in psychological therapy services based in primary and secondary care settings. *British Journal of Psychiatry*, **186**(3), 239–246. DOI:10.1192/bjp.186.3.239.
- Barkham M., Mellor-Clark J., Connell J., Cahill J. (2006) A core approach to practice-based evidence: a brief history of the origins and applications of the CORE-OM and CORE system. *Counselling and Psychotherapy Research*, **6**(1), 3–15. DOI:10.1080/14733140600581218.
- Burstein M., Georgiades K., Lamers F., Swanson S. A., Cui L., He J.-P., Avenevoli S., Merikangas K.R. (2012) Empirically derived subtypes of lifetime anxiety disorders: developmental and clinical correlates in U.S. adolescents. *Journal of Consulting and Clinical Psychology*, **80**(1), 102–115. DOI:10.1037/a0026069.
- Cahill J., Barkham M., Stiles W.B., Twigg E., Hardy G.E., Rees A., Evans C. (2006) Convergent validity of the CORE measures with measures of depression for clients in cognitive therapy for depression. *Journal of Counseling Psychology*, **53**(2), 253–259. DOI:10.1037/0022-0167.53.2.253.
- Central Statistics Office (2012) This is Ireland: Highlights from Census 2011, Part 1, Dublin: Stationary Office.
- Cheung G.W., Resnold R.B. (2002) Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, **9**(2), 233–255. DOI:10.1207/S15328007SEM0902_5.
- Connell J., Barkham M. (2007) CORE-10 User Manual, Version 1.1, Rugby: CORE System Trust & CORE Information Management Systems Ltd.
- Connell J., Barkham M., Stiles W.B., Twigg E., Singleton N., Evans O., Miles J.N.V. (2007) Distribution of CORE-OM scores in a general population, clinical cut-off points and comparison with the CIS-R. *British Journal of Psychiatry*, **190**(1), 69–74.
- Costello A.B., Osborne J.W. (2005) Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation*, **10**(7), 1–9.
- DeShon R.P. (2004) Measures are not invariant across groups without error variance homogeneity. *Psychology Science Quarterly*, **46**, 137–149.
- Drapeau A., Beaulieu-Prevost D., Marchand A., Boyer R., Preville M., Kairouz S. (2010) A life-course and time perspective on the construct validity of psychological distress in women and men. Measurement invariance of the K6 across gender. *BMC Medical Research Methodology*, **10**(1), 68. DOI:10.1186/1471-2288-10-68.
- Evans C., Connell J., Barkham M., Marshall C., Mellor-Clark J. (2003) Practice-based evidence: benchmarking NHS primary care counselling services at national and local levels. *Clinical Psychology & Psychotherapy*, **10**(6), 374–388. DOI:10.1002/cpp.384.
- Gorsuch R.L. (1997) Exploratory factor analysis: its role in item analysis. *Journal of Personality Assessment*, **68**(3), 532–560. DOI:10.1207/s15327752jpa6803_5.
- Gregorich S.E. (2006) Do self-report instruments allow meaningful comparisons across diverse population groups? Testing measurement invariance using the confirmatory factor analysis framework. *Medical Care*, **44**(11 Supplement 3), S78–S94. DOI:10.1097/01.mlr.0000245454.12228.8f.
- Hu L., Bentler P.M. (1999) Cutoff criteria for fit indices in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling*, **6**(1), 1–55.
- Jorm A.F., Windsor T.D., Dear K.B.G., Anstey K.J., Christensen H., Rodgers B. (2005) Age group differences in psychological distress: the role of psychosocial risk factors that vary with age. *Psychological Medicine*, **35**(9), 1253–1263. DOI:10.1017/S0033291705004976.
- Kessler R.C., Avenevoli S., Costello J., Green J.G., Gruber M.J., McLaughlin K.A., Petukhova M., Sampson N.A., Zaslavsky A.M., Merikangas K.R. (2012) Severity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication Adolescent Supplement. *Archives of General Psychiatry*, **69**(4), 381–389. DOI:10.1001/archgenpsychiatry.2011.1603.
- Kolenikov S., Angeles G. (2004) The Use of Discrete Data in PCA: Theory, Simulations, and Applications to Socioeconomic Indices, Chapel Hill, NC: MEASURE Evaluation.
- Kramer M.D., Krueger R.F., Hicks B.M. (2008) The role of internalizing and externalizing liability factors in accounting for gender differences in the prevalence of common psychopathological syndromes. *Psychological Medicine*, **38**(1), 51–61.
- Marshall K., Willoughby-Booth S. (2007) Modifying the clinical outcomes in routine evaluation measure for use with people who have a learning disability. *British Journal of Learning Disabilities*, **35**(2), 107–112. DOI:10.1111/j.1468-3156.2006.00422.x.
- McGorry P., Bates T., Birchwood M. (2013) Designing youth mental health services for the 21st century: examples from Australia, Ireland and the UK. *British Journal of Psychiatry*, **54**, 30–35. DOI:10.1192/bjp.bp.112.119214.
- McGorry P., Goldstone S., Parker A., Rickwood D., Hickie I. (2014) Cultures for mental health care of young people: an Australian blueprint for reform. *Lancet Psychiatry*, **1**(7), 559–568. DOI:10.1016/S2215-0366(14)00082-0.
- Mellor-Clark J., Connell J., Barkham M., Cummins P. (2001) Counselling outcomes in primary health care: a CORE system data profile. *European Journal of Psychotherapy & Counselling*, **4**(1), 65–86. DOI:10.1080/13642530110040118.
- Mental Health Reform (2014) Ethnic Minorities and Mental Health: A Position Paper, Dublin: Stationary Office.
- Merikangas K.R., He J.P., Burstein M., Swanson S.A., Avenevoli S., Cui L., Benjet C., Georgiades K., Swendsen J. (2010) Lifetime prevalence of mental disorders in U.S. adolescents: results from the National Comorbidity Survey Replication – Adolescent Supplement (NCS-A). *Journal of the American Academy of Child and Adolescent Psychiatry*, **49**(10), 980–989. DOI:10.1016/j.jaac.2010.05.017.

- Merikangas K.R., Nakamura E.F., Kessler R.C. (2009) Epidemiology of mental disorders in children and adolescents. *Dialogues in Clinical Neurosciences*, **11**(1), 7–20.
- O'Keeffe L., O'Reilly A., O'Brien G., Buckley R., Illback R. (2015) Description and outcome evaluation of Jigsaw: an emergent Irish mental health early intervention programme for young people. *Irish Journal of Psychological Medicine*, **32**(Special Issue 01), 71–77. DOI:10.1017/ipm.2014.86.
- O'Reilly A., Illback R., Peiper N., O'Keeffe L., Clayton R. (2015) Youth engagement with an emerging Irish mental health early intervention programme: participant characteristics and implications for service delivery. *Journal of Mental Health*, **24**(5), 285–288.
- Peiper N., Illback R.J., O'Reilly A., Clayton R. (2015) Latent class analysis of need descriptors within an Irish youth mental health early intervention program toward a typology of need. *Early Intervention in Psychiatry*, **11**(1), 7–20. DOI:10.1111/eip.12213.
- Rickwood D.J., Telford N.R., Parker A.G., Tanti C.J., McGorry P.D. (2014) headspace – Australia's innovation in youth mental health: who are the clients and why are they presenting? *Medical Journal of Australia*, **200**(2), 108–111.
- Sinclair A., Barkham M., Evans C., Connell J., Audin K. (2005) Rationale and development of a general population well-being measure: psychometric status of the GP-CORE in a student sample. *British Journal of Guidance and Counselling*, **33**(5), 153–173.
- Skre I., Friborg O., Elgarøy S., Evans C., Myklebust L.H., Lillevoll K., Sørgaard K., Hansen V. (2013). The factor structure and psychometric properties of the Clinical Outcomes in Routine Evaluation – Outcome Measure (CORE-OM) in Norwegian clinical and non-clinical samples. *BMC Psychiatry*, **13**(1), 99.
- StataCorp (2012) Stata Statistical Software: Release 12, TX, StataCorp LP: College Station.
- Twigg E., Barkham M., Bewick B.M., Mulhern B., Connell J., Cooper M. (2009) The young person's CORE: development of a brief outcome measure for young people. *Counselling and Psychotherapy Research*, **9**(3), 160–168. DOI:10.1080/14733140902979722.
- Twigg E., McInnes B. (2010) YP-CORE User Manual, Version 1.0, Rugby: CORE System Trust & CORE Information Management Systems Ltd.
- Vandenberg R.J., Lance C.E. (2000) A review and synthesis of the measurement invariance literature: suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, **3**(1), 4–70. DOI:10.1177/109442810031002.
- Wu A.D., Li Z., Zumbo B.D. (2007) Decoding the meaning of factorial invariance and updating the practice of multi-group confirmatory factor analysis: a demonstration with TIMSS data. *Practical Assessment, Research and Evaluation*, **12**(3), 1–26.